

An extremal problem in uniform distribution theory

Dedicated to Harald Niederreiter on the occasion of his 70th birthday

V. Baláž, M.R. Iacò, O. Strauch, S. Thonhauser, R.F. Tichy *

Abstract

In this paper we consider an optimization problem for Cesàro means of bivariate functions. We apply methods from uniform distribution theory, calculus of variations and ideas from the theory of optimal transport.

1 Introduction

In a series of papers J.G. van der Corput [26, 27] systematically investigated distribution functions of sequences of real numbers. More recently, the study of distribution functions was extended to multivariate functions by the Slovak school of O. Strauch and his coworkers; see [1, 2, 8, 23]. These investigations include the study of the set of all distribution functions of a given sequence and various optimization problems.

A particularly interesting problem is the study of extremal limits of two-dimensional sequences of the form

$$\frac{1}{N} \sum_{n=1}^N F(x_n, y_n), \quad N = 1, 2, \dots \quad (1)$$

*The authors were supported by the bilateral Austria-Slovakia travel grant “Uniform distribution, copulas and applications”. The second, forth and fifth authors are supported by the Austrian Science Fund (FWF) Project F5510 (part of the Special Research Program (SFB) “Quasi-Monte Carlo Methods: Theory and Applications”). The second author is also partially supported by the Austrian Science Fund (FWF): W1230, Doctoral Program “Discrete Mathematics”. The first and third authors are supported by VEGA Project 2/0149/14.

2010 *Mathematical Subject Classification.* 11K06, 60E05, 60A10, 49K27.

Keywords: Uniform distribution, copula, Monge-Kantorovich problem, dual problem, c -convex function, c -subdifferential.

where $(x_n)_{n \in \mathbb{N}}, (y_n)_{n \in \mathbb{N}}$ are uniformly distributed (for short u.d.) sequences in the unit interval and F is a given continuous function on $[0, 1]^2$; see [15].

Let us recall that a sequence $(x_n)_{n \in \mathbb{N}}$ of points in $[0, 1[$ is said to be u.d. if and only if

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \mathbf{1}_{[a, b[}(x_n) = b - a$$

for all intervals $[a, b[\subseteq [0, 1[$, where $\mathbf{1}_E$ denotes as usual the indicator function of the set E . We refer to [5, 12, 22] as general references on the subject.

A mapping T of the unit interval into itself is called uniform distribution preserving (for short u.d.p.) if the sequence $(T(x_n))_{n \in \mathbb{N}}$ is u.d. whenever $(x_n)_{n \in \mathbb{N}}$ is a u.d. sequence in $[0, 1[$.

These maps have been extensively studied (see e.g. [4, 16]), also in connection to variational problems [21] and extended to compact metric spaces [24]. They are particularly interesting for the purposes of this paper since they can be thought of as suitable measure preserving rearrangements of the unit interval, as we will see in the next section.

It turned out that the study of the asymptotic behaviour of mean values (1) is equivalent to find optimal bounds for Riemann-Stieltjes integrals of the form

$$\int_0^1 \int_0^1 F(x, y) dC(x, y), \quad (2)$$

where C is the asymptotic distribution function of the sequence $(x_n, y_n)_{n \in \mathbb{N}}$ and is usually referred to as copula (see [7]). More precisely, a 2-copula is a function $C: [0, 1]^2 \rightarrow [0, 1]$ satisfying the following properties: for every $x, y \in [0, 1]$

$$\begin{aligned} C(x, 0) &= C(0, y) = 0, \\ C(x, 1) &= x \text{ and } C(1, y) = y, \end{aligned}$$

and for every $x_1, x_2, y_1, y_2 \in [0, 1]$ with $x_2 \geq x_1$ and $y_2 \geq y_1$

$$C(x_2, y_2) - C(x_2, y_1) - C(x_1, y_2) + C(x_1, y_1) \geq 0.$$

An important property of copulas, which makes the expression in (2) meaningful and which can be derived from the above properties, is that every copula C induces a doubly stochastic measure γ_C (later on denoted by γ when there is no possibility of confusion) on the measurable space $([0, 1]^2, \mathcal{B})$, via the formula

$$\gamma([a, b] \times [c, d]) = C(b, d) - C(b, c) - C(a, d) + C(a, c).$$

Moreover, there is a one-to-one correspondence between copulas and doubly stochastic measures. For every copula C , the measure γ is doubly stochastic in the sense that for every Borel set $B \subset [0, 1]$, $\gamma([0, 1] \times B) = \gamma(B \times [0, 1]) = \lambda(B)$ where λ is the Lebesgue measure on $[0, 1]$. Conversely, for every doubly stochastic measure μ , there exists a copula C given by $C(u, v) = \mu([0, u] \times [0, v])$. Clearly, a probability measure on $([0, 1]^2, \mathcal{B})$ with uniform marginals is doubly stochastic. We refer to [6, 11, 14] for details.

With a slight abuse of notation we write $\int_0^1 \int_0^1 F(x, y) dC(x, y)$ for the integral with respect to the measure γ .

It should be remarked that copulas are very popular in applied probability. In particular, they are used in financial mathematics for modeling dependency structures among different kinds of risks; see [3, 13, 17, 19].

Notice, however, that in the calculation of these integrals one can also take advantage of the probabilistic interpretation of a copula, as the joint distribution function of a random vector with uniform marginals. More precisely, consider a random vector (X_1, X_2) and suppose that its marginals F_1 and F_2 are continuous. By applying the probability integral transformation to each component, the random vector

$$(U_1, U_2) = (F_1(X_1), F_2(X_2))$$

has uniform marginals. The copula of (X_1, X_2) is then defined as the joint cumulative distribution function (cdf) of (U_1, U_2) :

$$C(u_1, u_2) = P(U_1 \leq u_1, U_2 \leq u_2).$$

In this setting, the notion of optimal coupling can be stated. Let μ, ν be two probability measures on $(\mathbb{R}, \mathcal{B})$. Coupling μ and ν means constructing two random variables X_1, X_2 on \mathbb{R} in such a way that $X_1 \stackrel{d}{\sim} \mu$ and $X_2 \stackrel{d}{\sim} \nu$, with $\stackrel{d}{\sim}$ denoting equality in distribution. The couple (X_1, X_2) is called a coupling of (μ, ν) . Now, if we introduce a cost function $c(x, y)$ on $\mathbb{R} \times \mathbb{R}$, then the problem of finding

$$\inf \mathbb{E}(c(X_1, X_2))$$

where the pair (X_1, X_2) runs over all possible couplings of (μ, ν) is called c -optimal coupling or Monge-Kantorovich mass transportation problem (see

e.g. [18, 28]). Equivalently, this problem can be stated in terms of measures,

$$\inf \int c(x, y) dP(x, y),$$

where the infimum runs over all joint probability measures P on $\mathbb{R} \times \mathbb{R}$ with marginals μ and ν . For connections of extremal limits and copulas to the theory of optimal transport we refer to [10].

Through this paper we consider the equivalent sup problem

$$\sup \{ \mathbb{E}(c(X_1, X_2)) \mid X_1, X_2 \text{ couplings of } \mu, \nu \text{ with } P_{X_1} = \mu \text{ and } P_{X_2} = \nu \}. \quad (3)$$

Furthermore, we focus our attention on couplings between uniform distributions. In Section 4 we use this approach for solving a specific instance of maximizing the integral (2).

A useful criterion for checking the optimality of a candidate solution is based on the notion of c -convexity. A function $f : X \rightarrow \mathbb{R}$ is called c -convex if it has a representation $f(x) = \sup_y \{c(x, y) + a(y)\}$, for some function a . The associated c -subdifferential of f at x is then defined through:

$$\partial_c f(x) = \{y \mid f(z) - f(x) \geq c(z, y) - c(x, y) \ \forall z \in X\}$$

and $\partial_c f = \{(x, y) \in X \times Y \mid y \in \partial_c f(x)\}$.

Notice, see [18], that $y \in \partial_c f(x)$ if and only if $\exists a(=a(y)) \in \mathbb{R}$ such that

$$\psi_{y,a}(x) = c(x, y) + a(y) = f(x) \quad \text{and} \quad \psi_{y,a}(\xi) = c(\xi, y) + a(y) \leq f(\xi), \quad (4)$$

for every $\xi \in X$.

The dual problem of (3) is given by:

$$I(c) = \inf \left\{ \int h_1 d\mu + \int h_2 d\nu \mid c \leq h_1 + h_2, h_1 \in L^1(\mu) \text{ and } h_2 \in L^1(\nu) \right\},$$

and its study is the basis of the following theorem.

Theorem 1 (Th. 4.7 from [18]). *Let c be such that $c(x, y) \geq a(x) + b(y)$ for some $a \in L^1(\mu)$, $b \in L^1(\nu)$ and assume finiteness of $I(c)$. Then a pair (X_1, X_2) with $X_1 \stackrel{d}{\sim} \mu$, $X_2 \stackrel{d}{\sim} \nu$ is an optimal c -coupling between μ and ν if and only if*

$$(X_1, X_2) \in \partial_c f \text{ a.s.}$$

for some c -convex function f , equivalently, $X_2 \in \partial_c f(X_1)$ a.s.

2 Main results

As already pointed out in [7], the solution of problem (2) depends on the sign of the partial derivative $D_2 = \frac{\partial^2 F(x,y)}{\partial x \partial y}$. Special cases have been already studied in the literature, like those described in Fig.1 and Fig. 2. (see [7]). In particular, the upper and lower bounds for the first case are given precisely by the Fréchet-Hoeffding bounds, while in the second case the authors provide a criterion [7, Theorem 7] to find the corresponding extrema.

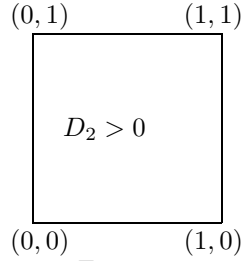


Figure 1.

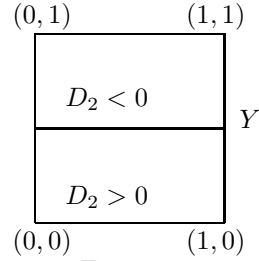


Figure 2.

In this paper we maximize (2) in the special situation described in Fig. 3 as a problem of optimal coupling (see [25]) and we provide a criterion for the instance of Fig. 4.

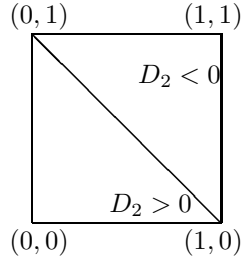


Figure 3.

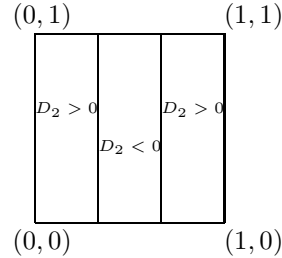


Figure 4.

We start by determining the copula which maximizes (2) when the sign of the second derivative changes as described in Fig. 4. We apply the following criterion [7, Theorem 7].

Theorem 2. *Let us assume that a copula $C(x,y)$ maximizes the integral $\int_0^1 \int_0^1 F(x,y) d\tilde{C}(x,y)$. Let $[X_1, X_2] \times [Y_1, Y_2]$ be an interval in $[0,1]^2$ such that*

$$C(X_2, Y_2) + C(X_1, Y_1) - C(X_1, Y_2) - C(X_2, Y_1) > 0 \quad (5)$$

and such that for every interior point (x,y) the mixed second derivative D_2 has constant sign. Then we have:

(i) if $D_2 > 0$, then

$$C(x, y) = \min(C(x, Y_2) + C(X_1, y) - C(X_1, Y_2), C(x, Y_1) + C(X_2, y) - C(X_2, Y_1)), \quad (6)$$

(ii) if $D_2 < 0$, then

$$C(x, y) = \max(C(x, Y_2) + C(X_2, y) - C(X_2, Y_2), C(x, Y_1) + C(X_1, y) - C(X_1, Y_1)), \quad (7)$$

for every $(x, y) \in [X_1, X_2] \times [Y_1, Y_2]$.

This result can be illustrated by the following pictures, where the symbols \oplus and \ominus in a corner mean that the value of C in that point is taken with positive and negative sign, respectively.

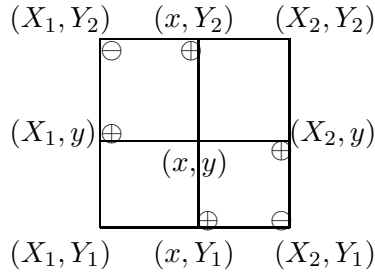


Figure 5.

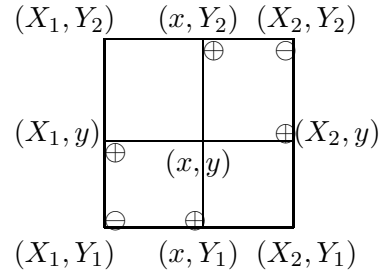


Figure 6.

In order to apply this criterion to the case described in Fig. 4 we divide the unit square $[0, 1]^2$ into $[0, x_1] \times [0, 1]$, $[x_1, x_2] \times [0, 1]$ and $[x_2, 1] \times [0, 1]$, as can be seen in Fig. 7.

Then, following the above statement, if $x \in (0, x_1) \cup (x_2, 1)$, we apply (6) in the same way as in Fig. 5 and if $x \in (x_1, x_2)$, we apply (7) as in Fig. 6.

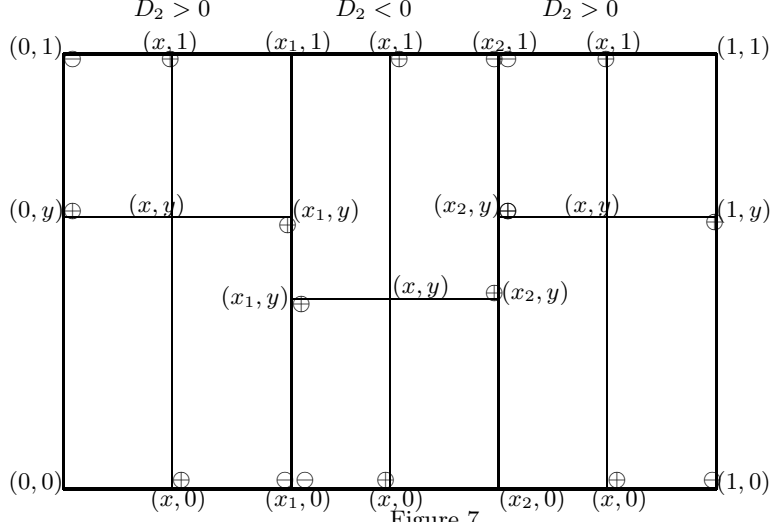


Figure 7.

Consequently, the following Theorem holds true.

Theorem 3. *Let $0 < x_1 < x_2 < 1$ and*

$$F(x, y) = \begin{cases} F_1(x, y) & \text{if } x \in (0, x_1), \frac{\partial^2 F_1(x, y)}{\partial x \partial y} > 0, \\ F_2(x, y) & \text{if } x \in (x_1, x_2), \frac{\partial^2 F_2(x, y)}{\partial x \partial y} < 0, \\ F_3(x, y) & \text{if } x \in (x_2, 1), \frac{\partial^2 F_3(x, y)}{\partial x \partial y} > 0. \end{cases} \quad (8)$$

Then the copula maximizing $\int_0^1 \int_0^1 F(x, y) d\tilde{C}(x, y)$ has the form

$$C(x, y) = \begin{cases} \min(x, h_1(y)) & \text{if } x \in [0, x_1], \\ \max(x + h_2(y) - x_2, h_1(y)) & \text{if } x \in [x_1, x_2], \\ \min(x - x_2 + h_2(y), y) & \text{if } x \in [x_2, 1], \end{cases} \quad (9)$$

where $h_1(y) = C(x_1, y)$, and $h_2(y) = C(x_2, y)$.

As we will see below, this result implies that in an ideal situation the problem is reduced to the determination of suitable functions h_1 and h_2 .

Before going on we need to determine $dC(x, y)$ for the special situation of (9). For this reason let us consider the rectangles

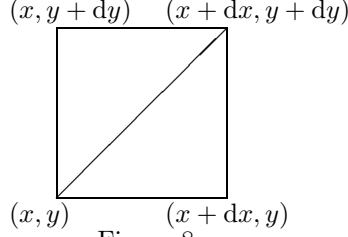


Figure 8.

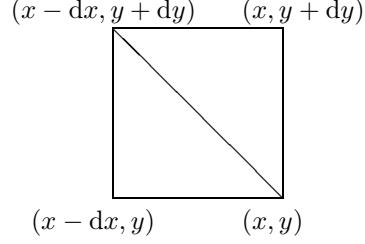


Figure 9.

and the - from the copula induced - measures which are defined by

$$\gamma_C(dx, dy) = C(x, y) + C(x + dx, y + dy) - C(x, y + dy) - C(x + dx, y) \quad (10)$$

and

$$\gamma_C(dx, dy) = C(x - dx, y) + C(x, y + dy) - C(x - dx, y + dy) - C(x, y), \quad (11)$$

where (dx, dy) stands for the infinitesimal rectangles from Fig. 8 and Fig. 9.

We consider the three regions in Fig. 7 where the second derivative changes sign separately.

(i) $x \in (0, x_1)$.

Then $x = h_1(y)$ and $C(x, y) = \min(x, h_1(y))$. Thus by (10)

$$\gamma_C(dx, dy) = h_1(y) + (h_1(y) + h_1'(y)dy) - h_1(y) - h_1(y) = h_1'(y)dy. \quad (12)$$

(ii) $x \in (x_1, x_2)$.

Then $x = x_2 - h_2(y) + h_1(y)$ and $C(x, y) = \max(x + h_2(y) - x_2, h_1(y))$.

Let us observe that

$$\begin{aligned} C(x, y + dy) &= \max(x + h_2(y + dy) - x_2, h_1(y + dy)) \\ &= \max(h_1(y) + h_2'(y)dy, h_1(y) + h_1'(y)dy) \\ &= h_1(y) + h_2'(y)dy, \end{aligned}$$

since for every (x, y) such that $x + h_2(y) - x_2 = h_1(y)$ we have $\frac{dx}{dy} + h_2'(y) = h_1'(y)$ and $\frac{dx}{dy} < 0$.

Similarly

$$\begin{aligned} C(x - dx, y) &= \max(x - dx + h_2(y) - x_2, h_1(y)) \\ &= \max(h_1(y) - dx, h_1(y)) = h_1(y), \end{aligned}$$

since $dx > 0$. Thus from (11) we have

$$\begin{aligned}\gamma_C(dx, dy) &= h_1(y) + (h_1(y) + h'_2(y)dy) - h_1(y) - h'_1(y) - h_1(y)dy \\ &= (h'_2(y) - h'_1(y))dy.\end{aligned}\tag{13}$$

(iii) $x \in (x_2, 1)$.

Then $x = x_2 - h_2(y) + y$ and $C(x, y) = \min(x - x_2 + h_2(y), y)$. Let us observe that

$$\begin{aligned}C(x + dx, y + dy) &= \min(x + dx - x_2 + h_2(y + dy), y + dy) \\ &= \min(y + dx + h'_2(y)dy, y + dy) = y + dy,\end{aligned}$$

since for every (x, y) such that $x - x_2 + h_2(y) = y$ we have $dx + h'_2(y)dy = dy$.

Moreover

$$\begin{aligned}C(x, y + dy) &= \min(x - x_2 + h_2(y + dy), y + dy) \\ &= \min(y + h'_2(y)dy, y + dy) = y + h'_2(y)dy,\end{aligned}$$

since $h'_2(y) \leq 1$ and

$$C(x + dx, y) = \min(x + dx - x_2 + h_2(y), y) = \min(y + dx, y) = y,$$

since $dx > 0$. Therefore with (10) we arrive at

$$\begin{aligned}\gamma_C(dx, dy) &= C(x, y) + C(x + dx, y + dy) - C(x, y + dy) - C(x + dx, y) \\ &= y + y + dy - (y + h'_2(y)dy) - y = (1 - h'_2(y))dy.\end{aligned}\tag{14}$$

Altogether the measure γ_C of the infinitesimal rectangles and hence $dC(x, y)$ is given by

$$\gamma_C(dx, dy) = \begin{cases} h'_1(y)dy & \text{if } x \in [0, x_1], x = h_1(y), \\ (h'_2(y) - h'_1(y))dy & \text{if } x \in [x_1, x_2], x = x_2 - h_2(y) + h_1(y), \\ (1 - h'_2(y))dy & \text{if } x \in [x_2, 1], x = x_2 - h_2(y) + y. \end{cases}\tag{15}$$

Our next step is to identify situations in which C is indeed a copula.

Theorem 4. *The function $C(x, y)$ defined by (9) is a copula if and only if*

(i) $h_1(y)$ and $h_2(y)$ are increasing;

(ii) $h_1(0) = 0$, $h_2(0) = 0$;

$$(iii) \quad h_1(1) = x_1, \quad h_2(1) = x_2;$$

$$(iv) \quad 0 \leq h_1(y) \leq h_2(y) \leq y;$$

$$(v) \quad 0 \leq h'_1(y) \leq h'_2(y) \leq 1.$$

Proof. The structure of the proof is as follows: we first prove the necessary condition by showing that if C is a copula, then properties (i) – (v) are satisfied. Then we exploit these properties to show that C is a copula.

Let $C(x, y)$ be a copula and $h_1(y) = C(x_1, y)$ and $h_2(y) = C(x_2, y)$.

Properties (i) – (iii) are straightforward. In order to prove (iv), let us consider the rectangle $[x, 1] \times [0, 1]$. Since $C(x, y)$ is a copula, we have

$$C(x, 0) + C(1, y_1) - C(x, y_1) - C(1, 0) = y_1 - C(x, y_1) \geq 0, \quad (16)$$

and thus $y_1 \geq C(x, y_1)$.

We proceed in a similar way to prove (v). Let us consider the rectangle $[x, 1] \times [y_1, y_2]$. For an arbitrary copula $C(x, y)$ we have

$$C(x, y_1) + C(1, y_2) - C(x, y_2) - C(1, y_1) = C(x, y_1) + y_2 - C(x, y_2) - y_1 \geq 0. \quad (17)$$

Then

$$y_2 - y_1 \geq C(x, y_2) - C(x, y_1)$$

and thus $C'(x, y) \leq 1$ a.e.. This implies $h'(y) \leq 1$ a.e. for $h(y) = C(x, y)$ (see also [14, Theorem 2.2.7.]). Furthermore

$$h'_1(y) \leq h'_2(y)$$

since for every (x, y) such that $x + h_2(y) - x_2 = h_1(y)$ with $x \in [x_1, x_2]$ we have $\frac{dx}{dy} + h'_2(y) = h'_1(y)$ and $\frac{dx}{dy} < 0$.

On the other hand, it follows from (v) and (15) that $\gamma_C(dx, dy)$ is nonnegative for every $(x, y) \in [0, 1]^2$ and by (9) that $C(x, 0) = C(0, y) = 0$. Thus $C(x, y)$ is a distribution function. We need to show that $C(x, 1) = x$ and $C(1, y) = y$ for every $(x, y) \in [0, 1]^2$. Indeed we have

$$C(x, 1) = \begin{cases} \min(x, h_1(1)) = \min(x, x_1) = x & \text{if } x \in [0, x_1], \\ \max(x + h_2(1) - x_2, h_1(1)) = \max(x, x_1) = x & \text{if } x \in [x_1, x_2], \\ \min(x - x_2 + h_2(1), 1) = \min(x, 1) = x & \text{if } x \in [x_2, 1]. \end{cases}$$

For $x = 1$ we need

$$C(1, y) = \min(1 - x_2 + h_2(y), y) = y. \quad (18)$$

Since $h_2(1) = x_2$, then (18) is equivalent

$$1 - y \geq h_2(1) - h_2(y), \quad (19)$$

which holds true, since $h_2(1) - h_2(y) = (1 - y)h'_2(y^*)$ for some $y^* \in (y, 1)$ and the derivative satisfies (v). \square

Theorem 4 implies the following bounds on candidate functions h_1 and h_2 .

$$\begin{aligned} \overline{h_1}(y) &= \begin{cases} y & \text{if } y \in [0, x_1], \\ x_1 & \text{if } y \in [x_1, 1], \end{cases} & \underline{h_1}(y) &= \begin{cases} 0 & \text{if } y \in [0, 1 - x_1], \\ y - (1 - x_1) & \text{if } y \in [1 - x_1, 1], \end{cases} \\ \overline{h_2}(y) &= \begin{cases} y & \text{if } y \in [0, x_2], \\ x_2 & \text{if } y \in [x_2, 1], \end{cases} & \underline{h_2}(y) &= \begin{cases} 0 & \text{if } y \in [0, 1 - x_2], \\ y - (1 - x_2) & \text{if } y \in [1 - x_2, 1], \end{cases} \end{aligned}$$

where

$$\underline{h_1}(y) \leq h_1(y) \leq \overline{h_1}(y), \quad \underline{h_2}(y) \leq h_2(y) \leq \overline{h_2}(y). \quad (20)$$

Now, we return to the integral (2).

Theorem 5. *Let us define a function G by*

$$\begin{aligned} G &:= G(y, h_1, h_2, h'_1, h'_2) = \\ &= F_1(h_1(y), y)h'_1(y) + F_2(x_2 - h_2(y) + h_1(y), y)(h'_2(y) - h'_1(y)) \\ &+ F_3(x_2 - h_2(y) + y, y)(1 - h'_2(y)). \end{aligned} \quad (21)$$

If h_1, h_2 maximize $\int_0^1 G dy$ and satisfy Theorem 4, then

$$\max_{C(x,y) \text{--} \text{copula}} \int_0^1 \int_0^1 F(x, y) dC(x, y) = \int_0^1 G dy. \quad (22)$$

If not, then we only have the following inequality

$$\max_{C(x,y) \in \mathcal{C}} \int_0^1 \int_0^1 F(x, y) dC(x, y) \leq \int_0^1 G dy. \quad (23)$$

The class \mathcal{C} is the set of all copulas of the form (9) with h_1, h_2 fulfilling the assumptions of Theorem 4.

Proof. Let F be a function defined on $[0, 1]^2$ such that $D_2 = \frac{\partial^2 F(x, y)}{\partial x \partial y}$ changes its sign as indicated in Fig. 4. Then the two-dimensional Riemann-Stieltjes integral of F with respect to the copula C defined in (9) is given as follows

$$\begin{aligned} \int_0^1 \int_0^1 F(x, y) dC(x, y) &= \int_0^1 F_1(h_1(y), y) h_1'(y) dy + \\ &+ \int_0^1 F_2(x_2 - h_2(y) + h_1(y), y) (h_2'(y) - h_1'(y)) dy \\ &+ \int_0^1 F_3(x_2 - h_2(y) + y, y) (1 - h_2'(y)) dy \quad (24) \\ &= \int_0^1 G dy . \quad (25) \end{aligned}$$

Since under the assumptions of Theorem 4., C is indeed a copula, the representation from Theorem 3. implies optimality. The second statement is obvious, since the class \mathcal{C} is a subset, due to additional restrictions, of candidate functions h_1, h_2 . \square

Remark 1. Note that to compute extremes of $\int_0^1 G(y, h_1, h_2, h_1', h_2') dy$ we can apply calculus of variations (cf [29, p. 33]). In particular, if (h_1, h_2) are extrema for the integral $\int_0^1 G(y, h_1, h_2, h_1', h_2') dy$, then (h_1, h_2) satisfy the Euler-Lagrange differential equations

$$\begin{aligned} \frac{\partial G}{\partial h_1} - \frac{d}{dy} \frac{\partial G}{\partial h_1'} &= 0 , \\ \frac{\partial G}{\partial h_2} - \frac{d}{dy} \frac{\partial G}{\partial h_2'} &= 0 . \end{aligned} \quad (26)$$

The solution (h_1, h_2) to (26) maximizes $\int_0^1 G(y, h_1, h_2, h_1', h_2') dy$ if

$$\frac{\partial^2 G}{\partial h_1' \partial h_1'} \leq 0, \quad \left| \begin{array}{cc} \frac{\partial^2 G}{\partial h_1' \partial h_1'} & \frac{\partial^2 G}{\partial h_1' \partial h_2'} \\ \frac{\partial^2 G}{\partial h_2' \partial h_1'} & \frac{\partial^2 G}{\partial h_2' \partial h_2'} \end{array} \right| \leq 0. \quad (27)$$

Remark 2. From the optimal copula with representation (9) and the properties of h_1 and h_2 from Theorem 4. we can derive the solution of the problem in the vocabulary of optimal couplings as well. Notice that for $x \in [0, x_1)$, (x, y) is mapped to $(h_1(y), y)$. According to Theorem 4. h_1 is monotone increasing and admits an inverse g_1 . For $x \in [x_1, x_2)$, we have (x, y) is mapped to $(x_2 - (h_2(y) - h_1(y)), y)$, where $x_2 - (h_2(y) - h_1(y))$ is monotone decreasing in y with inverse function g_2 . Finally, for $x \in [x_2, 1]$, we have

$(x, y) \mapsto (x_2 + y - h_2(y), y)$, with $x_2 + y - h_2(y)$ increasing in y and inverse g_3 . Therefore we can identify the optimal coupling $(U, \Gamma(U))$ for U uniformly distributed on $[0, 1]$ and

$$\Gamma(x) = \begin{cases} g_1(x), & x \in [0, x_1), \\ g_2(x), & x \in [x_1, x_2), \\ g_3(x), & x \in [x_2, 1]. \end{cases}$$

3 A piecewise linear cost function

Let

$$F(x, y) = \begin{cases} F_1(x, y) = \frac{x}{x_1}y & x \in (0, x_1), \\ F_2(x, y) = \frac{x_2 - x}{x_2 - x_1}y & x \in (x_1, x_2), \\ F_3(x, y) = \frac{x - x_2}{1 - x_2}y & x \in (x_2, 1), \end{cases}$$

with x -component as shown in Fig. 10.

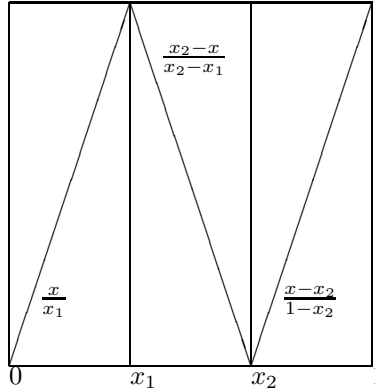


Figure 10.

Towards the construction of G from (21), we identify

$$\begin{aligned} F_1(h_1(y), y)h'_1(y) &= \frac{h_1}{x_1}yh'_1, \\ F_2(x_2 - h_2(y) + h_1(y), y)(h'_2(y) - h'_1(y)) &= \frac{h_2 - h_1}{x_2 - x_1}y(h'_2 - h'_1), \\ F_3(x_2 - h_2(y) + y, y)(1 - h'_2(y)) &= \frac{y - h_2}{1 - x_2}y(1 - h'_2), \end{aligned}$$

such that G takes the form

$$G = \frac{h_1}{x_1}yh'_1 + \frac{h_2 - h_1}{x_2 - x_1}y(h'_2 - h'_1) + \frac{y - h_2}{1 - x_2}y(1 - h'_2). \quad (28)$$

The associated Euler-Lagrange equations are given by

$$\frac{\partial G}{\partial h_1} - \frac{d}{dy} \frac{\partial G}{\partial h'_1} = \frac{h'_1}{x_1} y - \frac{h'_2 - h'_1}{x_2 - x_1} y - \frac{h_1}{x_1} + \frac{h_2 - h_1}{x_2 - x_1} = 0, \quad (29)$$

$$\frac{\partial G}{\partial h_2} - \frac{d}{dy} \frac{\partial G}{\partial h'_2} = \frac{h'_2 - h'_1}{x_2 - x_1} y - \frac{1 - h'_2}{1 - x_2} y - \frac{h_2 - h_1}{x_2 - x_1} + \frac{y - h_2}{1 - x_2} + \frac{y}{1 - x_2} = 0. \quad (30)$$

Now, adding (29) and (30) and multiplying the sum by $\frac{1-x_2}{y}$ we get

$$h'_1 \left(\frac{1 - x_2}{x_1} \right) + h'_2 = \frac{h_1}{y} \left(\frac{1 - x_2}{x_1} \right) + \frac{h_2}{y} - 1. \quad (31)$$

Multiplication of (29) with $\left(\frac{x_2 - x_1}{y} \right)$ gives

$$h'_1 \left(\frac{x_2 - x_1}{x_1} + 1 \right) - h'_2 = \frac{h_1}{y} \left(\frac{x_2 - x_1}{x_1} + 1 \right) - \frac{h_2}{y}. \quad (32)$$

Summing up (31) and (32) we find

$$h'_1 = \frac{h_1}{y} - x_1 \quad (33)$$

and hence

$$h'_2 = \frac{h_2}{y} - x_2. \quad (34)$$

The general solution $h(y)$ of the differential equation

$$h' = \frac{h}{y} - x, \quad (35)$$

has the form

$$h(y) = c y - x y \log y. \quad (36)$$

From the boundary conditions $h(1) = x$ and $h(0) = 0$, we find $h(y) = x y (1 - \log y)$ and $h'(y) = x(-\log y)$. Thus

$$h_1(y) = x_1 y (1 - \log y), \quad h_2(y) = x_2 y (1 - \log y), \quad (37)$$

which unfortunately do not satisfy condition (v) in Theorem 4. Finally, G is given through

$$G = y^2 \log y \left(\frac{x_2}{1 - x_2} \right) + y^2 (\log y)^2 \left(\frac{x_2}{1 - x_2} \right) + y^2$$

which yields the value

$$\int_0^1 G dy = \frac{x_2}{1-x_2} \left(-\frac{1}{27} \right) + \frac{1}{3}. \quad (38)$$

On the other hand, if $h_1(y) = x_1 y$ and $h_2(y) = x_2 y$, then $G = y^2$ and $\int_0^1 G dy = \frac{1}{3}$. Thus (37) does not maximize $\int_0^1 G dy$.

Note that for $F(x, y) = f(x)y$, with $f(x)$ uniform distribution preserving map (u.d.p.) we have

$$\max_{x_n, y_n \text{ are u.d.}} \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N f(x_n) y_n = \max_{C(x, y) \text{--copula}} \int_0^1 \int_0^1 F(x, y) dC(x, y), \quad (39)$$

$$\max_{x_n \text{--u.d., } \Phi \text{--u.d.p.}} \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N f(x_n) \Phi(x_n) = \max_{\Phi \text{--u.d.p.}} \int_0^1 f(x) \Phi(x) dx. \quad (40)$$

Then (40) is in general smaller than (39). By [7, Corollary 3] for u.d.p. $f(x)$ we have

$$\max_{\Phi \text{--u.d.p.}} \int_0^1 \int_0^1 f(x) \Phi(x) dx = \frac{1}{3}. \quad (41)$$

Therefore in our situation we have $\int_0^1 \int_0^1 F(x, y) dC(x, y) \geq \frac{1}{3}$.

Remark 3. The example points out the deficiencies of the variational formulation in the present context. When maximizing (2) it is essential to preserve the uniform distribution property of the marginals. In the formulation via the function (21), which led to the problem from the calculus of variations, this constraint is not present anymore and the maximization takes place over differentiable h_1, h_2 . On the other hand the optimal copula C in (9) with $h_1(y) = C(x_1, y)$ and $h_2(y) = C(x_2, y)$ does not enforce any smoothness properties, which implies that when solving (21) one does not necessarily get an upper bound for (2). The same reasoning suggests that when maximizing (21) over differentiable functions, which fulfill the conditions stated in Theorem 4, one in general derives a lower bound for (2).

4 A different approach using coupling

In this section we consider the example

$$F(x, y) = \sin(\pi(x + y))$$

and relate this problem to combinatorial optimization. In [9] an upper bound for

$$\int_{[0,1]^2} \sin(\pi(x+y))\gamma(dx, dy) . \quad (42)$$

was found by means of the Hungarian Algorithm and we will show that the copula found in [9] is indeed the one maximizing (42).

For dealing with this particular example one may utilize Theorem 1. from [25]. But since the proof of what we actually need is not given there, we re-state the following particular version of this result and give its proof.

Theorem 6. *Let μ, ν be the uniform distribution on $[0, 1]$ and the cost function $c(x, y) = \phi(x + y)$ with $\phi : [0, 2] \rightarrow \mathbb{R}$. In particular we assume that $\phi \in \mathcal{C}^2[0, 1]$ and that there is $k \in (0, 2)$ such that $\phi''(x) < 0$ for $x \in [0, k]$ and $\phi''(x) > 0$ for $x \in (k, 2]$. If $\beta \in (0, 1)$ denotes the solution to*

$$\phi(2\beta) - \phi(\beta) = \beta\phi'(\beta),$$

then

$$\Gamma(x) = \begin{cases} \beta - x, & x \in [0, \beta), \\ x, & x \in [\beta, 1], \end{cases}$$

induces by $(U, \Gamma(U))$ for some standard uniformly distributed U an optimal c -coupling between P and Q .

Proof. For the proof we proceed as proposed in [25] and [20]. Define the following functions:

$$\begin{aligned} f_1(x) &= x\phi'(\beta), \\ f_2(x) &= \frac{1}{2}(\phi(2x) - \phi(2\beta)) + \beta\phi'(\beta), \\ \psi^1(\xi) &= \phi(\beta - x + \xi) + x\phi'(\beta) - \phi(\beta), \\ \psi^2(\xi) &= \phi(x + \xi) - \frac{1}{2}\phi(2x) - \frac{1}{2}\phi(2\beta) + \beta\phi'(\beta). \end{aligned}$$

Furthermore set

$$f(x) = f_1(x)I_{[0,\beta)}(x) + f_2(x)I_{[\beta,1]}(x),$$

and put for $\xi \in [0, 1]$:

$$\psi_{\Gamma(x)}(\xi) = \begin{cases} \psi^1(\xi), & x \in [0, \beta), \\ \psi^2(\xi), & x \in [\beta, 1]. \end{cases}$$

Here $\psi_{\Gamma(x)}(\xi)$ plays the role of $\psi_{y,a}(\xi) = c(\xi, y) + a(y)$ with $y = \Gamma(x)$ in (4). Now the idea, following Theorem 1 and (4), is to show that $y = \Gamma(x)$ is in the c -subdifferential of $f(x)$ for all $x \in [0, 1]$ which implies optimality of this particular coupling and optimality of the distribution induced by $(U, \Gamma(U))$ for the transport problem. For the c -convexity of f and the subdifferential property we need to show:

$$\begin{aligned}\psi_{\Gamma(x)}(x) &= f(x) \quad \forall x \in [0, 1], \\ \psi_{\Gamma(x)}(\xi) &\leq f(\xi) \quad \forall \xi \in [0, 1].\end{aligned}$$

We start with showing that $\psi_{\Gamma(x)}(x) = f(x)$. For $x \in [0, \beta)$ we have that $\Gamma(x) = \beta - x$ and

$$\psi_{\Gamma(x)}(x) = \psi^1(x) = x\phi'(\beta) = f_1(x) = f(x).$$

For $x \in [\beta, 1]$ we have $\Gamma(x) = x$ and

$$\psi_{\Gamma(x)}(x) = \psi^2(x) = \frac{1}{2}(\phi(2x) - \phi(2\beta)) + \beta\phi'(\beta) = f_2(x) = f(x).$$

It remains to show $\psi_{\Gamma(x)}(\xi) \leq f(\xi)$ for all $(x, \xi) \in [0, 1] \times [0, 1]$.

At first we need some details on the location of β . From the definition of β and the mean value theorem we obtain $\exists \tau \in (\beta, 2\beta)$ with $\phi'(\tau) = \phi'(\beta)$. Since ϕ is concave on $[0, k]$ and convex on $(k, 2]$ we see that $\beta < k < \tau < 2\beta$. In the following we distinguish four situations.

- For $x \in [0, \beta)$ and $\xi \in [0, \beta)$ the functions are given by

$$\psi_{\Gamma(x)}(\xi) = \phi(\beta - x + \xi) + x\phi'(\beta) - \phi(\beta), \quad f(\xi) = \xi\phi'(\beta).$$

Therefore we need to check:

$$f(\xi) - \psi_{\Gamma(x)}(\xi) = (\xi - x)\phi'(\beta) - (\phi(\beta + \xi - x) - \phi(\beta)) =: F_1(x, \xi). \quad (43)$$

Obviously $F_1(x, \xi) = 0$ for $(0, \beta)$, $(0, 0)$ and (x, x) . In general we have by concavity on $[0, k]$ and the definition of β , $\phi(2\beta) = \phi(\beta) + \beta\phi'(\beta)$

$$\phi(\beta) + (\beta + \xi - x - \beta)\phi'(\beta) \geq \phi(\beta + \xi - x),$$

since $\beta + \xi - x \in [0, 2\beta]$, which proves that (43) is positive.

- For $x \in [0, \beta)$ and $\xi \in [\beta, 1]$ the functions are given by

$$\begin{aligned}\psi_{\Gamma(x)}(\xi) &= \phi(\beta - x + \xi) + x\phi'(\beta) - \phi(\beta), \\ f(\xi) &= \frac{1}{2}(\phi(2\xi) - \phi(2\beta)) + \beta\phi'(\beta).\end{aligned}$$

We need the following to be positive:

$$\begin{aligned}& \frac{1}{2}(\phi(2\xi) - \phi(2\beta)) + \beta\phi'(\beta) - \phi(\beta + \xi - x) - x\phi'(\beta) + \phi(\beta) \\ &= \frac{1}{2}\phi(2\xi) + \frac{1}{2}\phi(2\beta) - x\phi'(\beta) - \phi(\beta + \xi - x) \\ &\geq \phi(\xi + \beta) - x\phi'(\beta) - \phi(\beta + \xi - x) =: F_2(x, \xi),\end{aligned}$$

where the equality above follows from the definition of β and the inequality follows from convexity since $k < 2\beta \leq 2\xi$. Observe $F_2(x, \xi) = 0$ for $(0, \xi)$ and (β, β) . To prove positivity the strategy is as follows, firstly show $F_2(x, \xi = \beta) \geq 0$ for all $x \in [0, \beta)$ and secondly show that $\frac{\partial}{\partial \xi} F_2(x, \xi) \geq 0$ for all $(x, \xi) \in [0, \beta) \times (\beta, 1)$.

Look at

$$\frac{\partial}{\partial x} F_2(x, \xi = \beta) = -\phi'(\beta) + \phi'(2\beta - x)$$

which is zero in $(0, \beta)$ exactly if $x = \hat{x} = 2\beta - \tau < \beta$. Since $0 < \phi''(\tau) = -\frac{\partial^2}{\partial x^2} F_2(x, \xi = \beta)|_{x=\hat{x}}$ we have that

$$F_2(\hat{x}, \xi = \beta) = \phi(2\beta) - (2\beta - \tau)\phi'(\tau) - \phi(\tau) > 0$$

is a maximum, it is positive by convexity ($k < \tau < 2\beta$). Thus, $F_2(x, \xi = \beta) \geq 0$ for $x \in [0, \beta)$. Now we can deal with checking the interior,

$$\frac{\partial}{\partial \xi} F_2(x, \xi) = \phi'(\beta + \xi) - \phi'(\beta + \xi - x).$$

Suppose $\beta + \xi - x \geq k$, then by convexity $\phi'(\beta + \xi) \geq \phi'(\beta + \xi - x)$. On the other hand if $\beta + \xi - x < k$ we have $\phi'(\beta + \xi - x) < \phi'(\beta) = \phi'(\tau) \leq \phi'(2\beta) \leq \phi'(\beta + \xi)$ since $2\beta \leq \beta + \xi$.

- Consider $x \in [\beta, 1]$ and $\xi \in [0, \beta)$, here

$$\psi_{\Gamma(x)}(\xi) = \phi(x + \xi) - \frac{1}{2}\phi(2x) - \frac{1}{2}\phi(2\beta) + \beta\phi'(\beta), \quad f(\xi) = f_1(\xi) = \xi\phi'(\beta),$$

and we need

$$\begin{aligned} & \xi\phi'(\beta) - \phi(x + \xi) + \frac{1}{2}\phi(2x) + \frac{1}{2}\phi(2\beta) - \beta\phi'(\beta) \\ & \geq \phi(x + \beta) - \phi(x + \xi) - (\beta - \xi)\phi'(\beta) =: F_3(x, \xi) \end{aligned}$$

to be positive, the inequality stems from convexity since $2x \geq 2\beta > k$. We proceed as before. $F_3(x, \xi) = 0$ for (x, β) and $(\beta, 0)$. At first fix $x = \beta$, $F_3(x = \beta, y) = 0$ if $y = 0$ and $y = \beta$. In between we study

$$\frac{\partial}{\partial \xi} F_3(x = \beta, \xi) = -\phi'(\beta + \xi) + \phi'(\beta),$$

which is zero in $(0, \beta)$ exactly if $\xi = \hat{\xi} = \tau - \beta$. Again due to convexity of ϕ on $(k, 2]$ we have a maximum in $\hat{\xi}$ and

$$F_3(x = \beta, \hat{\xi}) = \phi(2\beta) - \phi(\tau) - (2\beta - \tau)\phi'(\tau) > 0,$$

such that we have $F_3(x = \beta, \xi) \geq 0$ for $\xi \in [0, \beta)$. On the interior we show that

$$\frac{\partial}{\partial x} F_3(x, \xi) = \phi'(\beta + x) - \phi'(x + \xi) \geq 0.$$

If $x + \xi \geq k$ we have from convexity $\phi'(\beta + x) \geq \phi'(x + \xi)$. If $x + \xi < k$ we have $\beta \leq x + \xi < k \leq \tau < 2\beta \leq \beta + x$ and therefore

$$\phi'(x + \xi) \leq \phi'(\beta) = \phi'(\tau) \leq \phi'(2\beta) \leq \phi'(\beta + x).$$

- Let $x \in [\beta, 1]$ and $\xi \in [\beta, 1]$. Here

$$\begin{aligned} \psi_{\Gamma(x)}(\xi) &= \phi(x + \xi) - \frac{1}{2}\phi(2x) - \frac{1}{2}\phi(2\beta) + \beta\phi'(\beta), \\ f(\xi) &= \frac{1}{2}(\phi(2\xi) - \phi(2\beta)) + \beta\phi'(\beta), \end{aligned}$$

such that

$$f(\xi) - \psi_{\Gamma(x)}(\xi) = \frac{1}{2}\phi(2x) + \frac{1}{2}\phi(2\xi) - \phi(x + \xi) \geq 0,$$

is fulfilled by convexity since $x + \xi, 2x, 2\xi \geq 2\beta > k$.

We can conclude that $\psi_{\Gamma(x)}(\xi) \leq f(\xi) \quad \forall (x, \xi) \in [0, 1] \times [0, 1]$, which according to Theorem 1 shows that the vector $(U, \Gamma(U))$ for $U \stackrel{d}{\sim} \mathcal{U}([0, 1])$ yields an optimal coupling. \square

Remark 4. If $\beta > 1$ then it can be shown as in the first step of the above proof that $(U, 1-U)$ yields the optimal coupling. Loosely speaking one could say that the concave behaviour dominates the convex one.

Now we are prepared to answer the sine question. Setting $\phi(z) = \sin(\pi z)$ and $k = 1$ we immediately get:

Corollary 1. *For $c(x, y) = \sin(\pi(x + y))$ we have that the distribution of the vector $(U, \Gamma(U))$ for $U \sim \mathcal{U}([0, 1])$ with*

$$\Gamma(x) = \begin{cases} \beta - x, & x \in [0, \beta), \\ x, & x \in [\beta, 1], \end{cases}$$

and $\beta = 0.7541996008265638 \approx 0.7542$ (cf. [9]) which solves

$$\sin(2\pi\beta) - \sin(\pi\beta) = \beta\pi \cos(\pi\beta), \quad (44)$$

is maximizing

$$\int_{[0,1]^2} \sin(\pi(x + y)) dC(x, y)$$

in the set of all bivariate distributions C with uniform marginals, i.e., in the set of all copulas.

Remark 5. In this situation equation (44) meets the first order condition when looking at couplings of the form $(U, \Gamma^\alpha(U))$ with

$$\Gamma^\alpha(x) = \begin{cases} \alpha - x, & x \in [0, \alpha), \\ x, & x \in [\alpha, 1], \end{cases}$$

or explicitly maximizing $(c(x, y) = \sin(\pi(x + y)))$

$$H(\alpha) := \int_0^\alpha c(x, \alpha - x) dx + \int_\alpha^1 c(x, x) dx.$$

This is conjectured in [9] but with inaccurate $\alpha = \frac{3}{4}$.

Acknowledgments

M.R. Iacò and R.F. Tichy would like to thank the Slovak colleagues for their hospitality during the visit in Bratislava in November 2014.

References

- [1] V. Baláž, L. Mišík, O. Strauch, and J. T. Tóth. Distribution functions of ratio sequences, III. *Publ. Math. Debrecen*, 82(3-4):511–529, 2013.
- [2] V. Baláž, L. Mišík, O. Strauch, and J. T. Tóth. Distribution functions of ratio sequences, IV. *Period. Math. Hungar.*, 66(1):1–22, 2013.
- [3] M. Beiglböck, P. Henry-Labordère, and F. Penkner. Model-independent bounds for option prices—a mass transport approach. *Finance Stoch.*, 17(3):477–501, 2013.
- [4] W. Bosch. Functions that preserve the uniform distribution of sequences. *Trans. Amer. Math. Soc.*, 307(1):143–152, 1988.
- [5] M. Drmota and R. F. Tichy. *Sequences, discrepancies and applications*, volume 1651 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1997.
- [6] F. Durante and C. Sempi. *Principles of copula theory*. CRC/Chapman & Hall, London, 2015.
- [7] J. Fialová and O. Strauch. On two-dimensional sequences composed by one-dimensional uniformly distributed sequences. *Unif. Distrib. Theory*, 6(1):101–125, 2011.
- [8] G. Grekos and O. Strauch. Distribution functions of ratio sequences. II. *Unif. Distrib. Theory*, 2(1):53–77, 2007.
- [9] M. Hofer and M. R. Iacò. Optimal bounds for integrals with respect to copulas and applications. *Journal of Optimization Theory and Applications*, 161(3):999–1011, 2014.
- [10] M. R. Iacò, , R. F. Tichy, and S. Thonhauser. Distribution functions, extremal limits and optimal transport. *Preprint*, 2015.
- [11] P. Jaworski, F. Durante, W. Härdle, and T. Rychlik, editors. *Copula theory and its applications*, volume 198 of *Lecture Notes in Statistics—Proceedings*. Springer, Heidelberg, 2010.
- [12] L. Kuipers and H. Niederreiter. *Uniform distribution of sequences*. Wiley-Interscience [John Wiley & Sons], New York-London-Sydney, 1974. Pure and Applied Mathematics.
- [13] A. J. McNeil, R. Frey, and P. Embrechts. *Quantitative risk management*. Princeton Series in Finance. Princeton University Press, Princeton, NJ, 2005. Concepts, techniques and tools.
- [14] R. B. Nelsen. *An introduction to copulas*. Springer Series in Statistics. Springer, New York, second edition, 2006.
- [15] F. Pillichshammer and S. Steinerberger. Average distance between consecutive points of uniformly distributed sequences. *Unif. Distrib. Theory*, 4(1):51–67, 2009.
- [16] Š. Porubský, T. Šalát, and O. Strauch. Transformations that preserve uniform distribution. *Acta Arith.*, 49(5):459–479, 1988.
- [17] G. Puccetti and L. Rüschendorf. Sharp bounds for sums of dependent risks. *J. Appl. Probab.*, 50(1):42–53, 2013.
- [18] L. Rüschendorf. Monge-Kantorovich transportation problem and optimal couplings. *Jahresber. Deutsch. Math.-Verein.*, 109(3):113–137, 2007.

- [19] L. Rüschendorf. *Mathematical risk analysis*. Springer Series in Operations Research and Financial Engineering. Springer, Heidelberg, 2013. Dependence, risk bounds, optimal allocations and portfolios.
- [20] L. Rüschendorf and L. Uckelmann. Numerical and analytical results for the transportation problem of Monge-Kantorovich. *Metrika*, 51(3):245–258 (electronic), 2000.
- [21] S. Steinerberger. Uniform distribution preserving mappings and variational problems. *Unif. Distrib. Theory*, 4(1):117–145, 2009.
- [22] O. Strauch and Š. Porubský. *Distribution of sequences: a sampler*, volume 1 of *Schriftenreihe der Slowakischen Akademie der Wissenschaften [Series of the Slovak Academy of Sciences]*. Peter Lang, Frankfurt am Main, 2005.
- [23] O. Strauch and J. T. Tóth. Distribution functions of ratio sequences. *Publ. Math. Debrecen*, 58(4):751–778, 2001.
- [24] R. F. Tichy and R. Winkler. Uniform distribution preserving mappings. *Acta Arith.*, 60(2):177–189, 1991.
- [25] L. Uckelmann. Optimal couplings between one-dimensional distributions. In *Distributions with given marginals and moment problems (Prague, 1996)*, pages 275–281. Kluwer Acad. Publ., Dordrecht, 1997.
- [26] J. G. van der Corput. Verteilungsfunktionen I-II. *Proc. Akad. Amsterdam*, 38:813–821, 1058–1066, 1935.
- [27] J. G. van der Corput. Verteilungsfunktionen III-VIII. *Proc. Akad. Amsterdam*, 39:10–19, 19–26, 149–153, 339–344, 489–494, 579–590, 1936.
- [28] C. Villani. *Optimal transport*, volume 338 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2009. Old and new.
- [29] R. Weinstock. *Calculus of variations*. Dover Publications, Inc., New York, 1974. With applications to physics and engineering, Reprint of the 1952 edition.